IJLST

ISSN: 0974-5335
IJLST (2013), 6(2):9-16

Visualization of Overfitted Classifiers with Application to the Computer-Aided Diagnosis of Breast Cancer from Dynamic Contrast-Enhanced Magnetic Resonance Imaging

Jacob Levman[1], Anne Martel[2]

1 Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, United Kingdom
jacob.levman@eng.ox.ac.uk
1 & 2 Imaging Research, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Canada

ABSTRACT:
This research paper presents a classifier visualization technique as a method for identifying overfitted supervised learning classification functions and is demonstrated on breast cancer data from magnetic resonance imaging examinations. A classifier is a supervised learning algorithm that assigns a class (for example cancerous or benign tissue) to an unknown sample based on a labeled set of training data. An overfitted classifier is one that is too closely tuned to the training data that it was provided with, causing it to potentially perform poorly on other datasets. Overfitted classification functions are a major concern of pattern recognition researchers as they can yield extremely enticing test results while simultaneously providing an unreliable test. Randomized statistical testing can avoid overfitted solutions, however, the implementation of statistical testing is complex and prone to error. This study presents a plotting technique meant to visually illustrate a classification function making it easier to identify overfitted solutions. Example overfitted solutions are demonstrated for the k-nn density classifier as well as the established and high performing support vector machine. By allowing the researcher to easily visually identify overfitted solutions, the technique can act as an aid in producing robust computer-aided diagnosis systems that avoid the problems associated with overfitted classifiers.

Keywords: Visualization, overfitting, classification, supervised learning, breast cancer imaging, computer-aided diagnosis

INTRODUCTION

Computer-aided diagnosis systems are one of the largest applications of pattern recognition in medical imaging. Typical computer-aided diagnostic systems contain a critical step whereby the algorithm needs to decide if a given sample is normal or abnormal. The computer-aided diagnosis researcher wants this decision function (also known as a classification function or supervised learning technique) to be as robust as possible in order to ensure that the system developed will function well when exposed to new independent sets of medical imaging data that it wasn't trained on. One of the largest problems for the computer-aided diagnosis researcher to avoid is creating an overfitted classification function. An overfitted classification function is one that is overly tuned to the training data that it has been provided with and as such will probably not perform well when applied to unseen data acquired at a different imaging centre. Overfitting is a challenging subject to address as there is no gold standard quantitative definition for overfitting. Qualitatively, overfitting can be described as a classification function that is too closely tuned to the data it has been provided with, this often manifests as a classification function whose decision boundary changes dramatically across the data space (as opposed to a robust and relatively smooth non-overfitted function).

Computer-aided diagnosis publications don't normally include the visualization of classification solutions in order to help identify overfitting. Typically, overfitted classification functions are simply avoided as best as possible by performing extensive randomized validation [1] (for example leave-one-out, k-fold cross validation, randomized trials and bootstrapping [2]). These validation techniques involve separating our dataset into two groups, one for training and one for testing and then computing a metric which indicates how well the classifier is performing (such as the receiver operating characteristic curve area [3], the test's overall accuracy or some combination of the test's sensitivity and specificity). This process of dividing our data into training and testing groups is performed randomly and repeated many times leading to a distribution of test metric values. These distributions are used as the main method for evaluating the supervised learning technique. If the validation was performed well then the selected best performing classification function (ideally) is not overfitted. In practice this is not necessarily the case as the process of proper validation is complicated, researchers don't necessarily choose the best validation options and even if they do, a small mistake in their validation software can still yield overfitted solutions. Furthermore, some validation techniques such as leave-one-out validation and k-fold cross validation are typically trained on an extremely large percentage of the total samples available. When only a few samples are left out for testing, there is still a possibility of producing an overfitted solution.

This research study presents an existing visualization technique [1] as a novel method for qualitatively identifying overfitted classification solutions. This visualization technique can act as an adjunct to

existing validation techniques to help the computer-aided diagnosis researcher produce robust classification solutions. It should further be noted that the proposed method for visually detecting overfitted solutions has applications outside of computer-aided diagnosis as it can be applied to any classification function. However, in this paper an emphasis is made on this technique's ability to assist in avoiding overfitted computer-aided diagnosis decision functions with application to breast cancer detection from magnetic resonance imaging.

Computer-aided diagnosis systems have been developed to assist in the identification of a wide variety of pathologies from many types of medical imaging modalities. In this study we are focusing on demonstrating a high dimensional visualization technique's ability to detect overfitted classification solutions for breast cancer diagnosis from magnetic resonance imaging. Computer-aided diagnosis for breast cancer from magnetic resonance imaging is an active area of research, motivated by the known effect that there is a high amount of variability between trained radiologists in their interpretation of these breast examinations [4]. Furthermore, breast computer-aided diagnosis systems assist in processing the large amounts of data acquired during a magnetic resonance breast examination (hundreds of images compared with just 4 images from traditional x-ray mammogram based breast cancer screening). It is also known that cancerous lesions can be caught by magnetic resonance imaging when they are as small as just 2 to 3 mm across [5]. Having a radiologist catch a tiny 2-3 mm tumour in a dataset of hundreds of images can be extremely challenging (a 2-3 mm tumour may have as few as 12 visible voxels out of an exam containing 1.8 million voxels [5]) and is the type of challenge that an automated computer should be able to help with.

A computer-aided diagnosis system for breast cancer detection from magnetic resonance imaging makes a critical decision regarding whether a tissue is cancerous or non-cancerous. Typically, a computer-aided diagnosis system acquires measurements directly from the magnetic resonance breast examination and needs to combine that set of measurements in order to predict whether the collection of measurements represent cancer or non-cancer at any given location in the image. This decision is based on previous samples that were collected and provided to this classification program with known labels (cancer or benign). This classification step can be performed by many different technologies. In this study we demonstrate that existing classification techniques that have been used in the computer-aided diagnosis of breast cancer from magnetic resonance imaging such as the k-nn density classifier [6] and the support vector machine [1, 6] can all produce overfitted solutions.

The visualization technique presented can be used to identify overfitted solutions and generally to evaluate classifier behaviour. The plotting technique can also act as a simple sanity-check to make sure that the classifier conforms to a sensible robust solution. While this paper presents a novel approach to visualizing overfitted classification functions, some related work in visualization has been conducted on this topic. Nattkemper and Wismuller have researched the use of self organizing maps, an implementation of the artificial neural network for tumour feature visualization from breast MR data [7]. Komura *et al*. proposed the use of multidimensional support vector machines for the visualization of high dimensional data sets and demonstrated the approach with gene expression data [8]. Somorjai and Dolenko have proposed the visualization of high dimensional data onto a plane called the relative distance map [9]. These studies were focused on visualization, not on overfitting. Some related work on overfitting has been performed [10-18] and a few methods have been developed to assist in avoiding creating overfitted classifiers [19-23]. This paper demonstrates that a visual approach that has previously been published as a simple plotting method for evaluating classifiers [1] can be used to identify overfitted multidimensional classifiers.

MATERIALS AND METHODS

2.1 Breast MRI Data

The screening protocol used to acquire the magnetic resonance imaging based breast data for this study is as follows. Simultaneous bilateral magnetic resonance imaging was performed using a 1.5T magnet (General Electric Signa, version 11.4). Sagittal images were obtained with a phased-array coil arrangement using a dual slab interleaved bilateral imaging method [24]. This provided 3D volume data over each breast obtained with an RF spoiled gradient recalled sequence (SPGR, scan parameters: TR/TE/angle=18.4/4.3/30^{o}, 256x256x32 voxels, Field of View: 18x18x6-8cm). Imaging was performed before and after a bolus injection of 0.1 mmol/kg of Gd-DTPA. Each bilateral acquisition was obtained in 2 minutes and 48 seconds. The spatial resolution was 0.7 x 0.7 mm and the slice thickness was 2 to 3 mm.

Ethics approval was obtained from the institution in which this research was conducted (Sunnybrook Health Sciences Centre). Patients enrolled in research based screening consented in writing. A total of 259 DCE-MRI breast lesions from high risk patients were obtained yielding lesions pathologically proven to be malignant (51 cases) or benign (208 cases). Ground truth is based on the diagnosis of the histopathologist, who analyzes the tissue biopsies. Surgical biopsy of lesions was performed under magnetic resonance

imaging based guidance [25]. In cases where a patient with a suspicious lesion did not receive a biopsy but returned to screening for greater than one year without observed changes to the lesion, a benign diagnosis is accepted.

Image registration is the process of aligning images that vary in position over time. This is performed in order to compensate for any patient motion that takes place during the examination. For this study we have used a three-dimensional non-rigid registration technique for magnetic resonance breast images [26].

For each radiologically identified suspicious lesion, the maximally enhancing voxel's signal intensity time measurements were taken as the main feature vector (the set of measurements that a classifier will combine). These post contrast injection signal intensity measurements were normalized by the precontrast injection signal intensity, forming a four dimensional relative signal intensity feature measurement for each radiologically suspicious lesion.

2.2 Computer-Aided Diagnosis of Breast Cancer from Magnetic Resonance Imaging

A computer-aided diagnosis system for detecting breast cancer from magnetic resonance imaging examinations can employ one of many different mechanisms for making the final decision of whether a given sample is representative of cancer.

In this study we demonstrate that existing classification techniques that have been used for the computer-aided diagnosis of breast cancer from magnetic resonance imaging (and are widely used elsewhere) can produce overfitted solutions and that the presented visualization technique can assist the researcher in identifying them. We demonstrate the popular method known as the k-nn density classifier [6] as well as the established and high performing support vector machine [1, 6]. For the support vector machine the publicly available LIBSVM library was used [27], whereas the k-nn density classifier was developed in the research lab.

The computer-aided diagnosis researcher can choose between one of any of the supervised learning applications available in order to perform the critical step of deciding if a set of measurements represent cancer or non-cancerous tissues. In this study, the range of plausible parameter settings was explored for k in the k-nn density classifier and γ for the radial basis function based support vector machine.

This study demonstrates a plotting technique which is interpreted qualitatively, as such examples qualitatively described as both robust and overfitted from the support vector machine classifier are provided along with an overfitted example of the k-nn density classifier.

2.3 Proposed Method for Identifying Overfitted Classification Solutions

Overfitting is a major concern of computer-aided diagnosis researchers and it was thought that being able to visualize a classification function with respect to our dataset could be a beneficial means of identifying overfitted and generally undesirable classification solutions. Unfortunately, since our data is four dimensional in nature, visualization can be challenging. Here we are presenting a method for visualizing an overfitted classification function that separates data of a high dimensional nature. Although this technique has been introduced previously as a means of visually assessing a classification function [1], in this paper we are presenting the technique as a visual method for identifying overfitted multi-dimensional classification functions.

The first step of this technique is to project our high dimensional data into an easily visible two-dimensional space. This can be accomplished by a number of data projection techniques but we have elected to use the well-known principal components analysis. Principal components analysis is a mathematical technique for rotating high dimensional data such that the resultant orthogonal axes are aligned to the maximum variance in the data set. This rotation is performed and the two principal components with the highest corresponding eigenvalues are selected for viewing (referred to as the first and second principal components). Each rotated input vector is plotted on a grid representing these first two principal components (see figure 1 upper pane for an example plot).

The second step of the technique is to sample this visualized principal component space across the range of first and second principal component values that our dataset occupies. Thus we sample from the minimum principal component value in our data to the maximum principal component value, however, if necessary, a larger sampling area can be used. In the case of the plots presented in the results of this paper, principal component space was sampled 400 times along each of the first two principal components. Each of these sampled points is then reverse rotated back to the input space and the resultant signal-intensity time vector is compared against any given classifier prediction technique. This generates a binary image (in our case 400x400) where the regions represent how the given classifier predicts classes in principal component space; an example binary image is provided in figure 1, middle pane. This binary image is processed through an edge detector (we used the Sobel edge detector) and the resultant edge is plotted in the original principal component space on top of the rotated training data our classifier is using. An example plot is provided in figure 1 bottom pane. The resultant contour line represents the border between different predictions for a given classifier in principal component space.

The percentage of total data variance displayed in the resultant two-dimensional plot can be computed as the sum of the first two principal component eigenvalues (the two largest eigenvalues) divided by the sum of all of the eigenvalues. The classification function can also be evaluated in the lower principal components by extending this technique to a set of pair-wise plots of principal component space (plotting not just the first principal component against the second principal component, but also the second principal component against the third principal component and so on). It should be noted that any method of projecting high-dimensional data into a two-dimensional space inherently involves loss of some of the information available. It is impossible to guarantee that any given projection is the ideal two-dimensional plot on which to evaluate a given classification function. The quality of using this technique as a classification evaluation mechanism is dependent on how much of the total data variance is captured in the principal component rotation process.

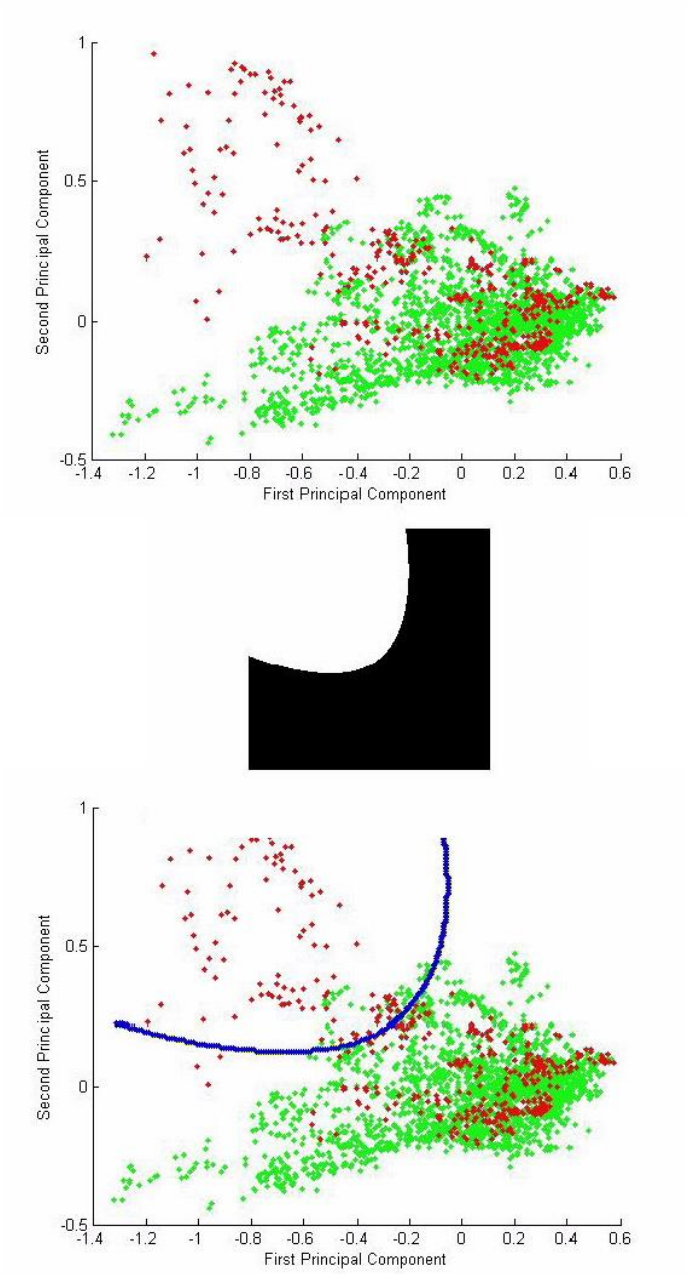

Fig. 1 shows an example plot of breast MRI data in principal component space (top pane) where the red dots mark malignant samples and the green dots mark benign samples. An example binary image produced from reverse rotation and classification is provided (middle pane) and the final high dimensional classifier visualized plot in principal component space is provided in the bottom pane.

RESULTS

Plots produced by the proposed technique are provided. Figure 2 demonstrates an overfitted support vector machine classifier ($\gamma = e^{8}$). Figure 3 demonstrates an overfitted k-nn density classifier (k=1). Figure 4 illustrates a robust support vector machine classifier ($\gamma = e^{-2}$). Finally, figure 5 provides an example magnetic resonance image of the breast (top pane) along with computer-aided diagnosis results of a support vector machine classifier ($\gamma = e^{-2}$, center pane) and computer-aided diagnosis results of an overfitted support vector machine classifier ($\gamma = e^{8}$, bottom pane).

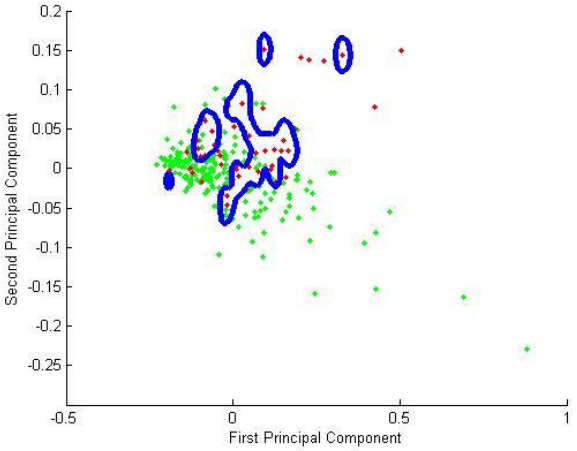

Fig. 2 Principal component projection plot of high-dimensional breast cancer data with an overfitted radial basis function based support vector machine ($\gamma = e^{8}$) projected onto the data (blue contour lines).

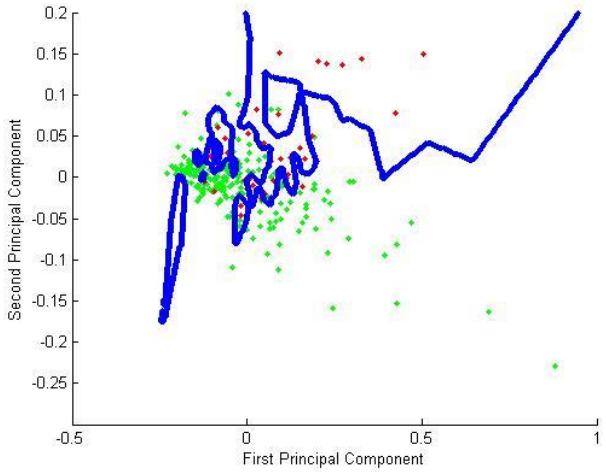

Fig. 3. Principal component projection plot of high-dimensional breast cancer data with an overfitted k-nn density classifier (k=1) projected onto the data (blue contour lines).

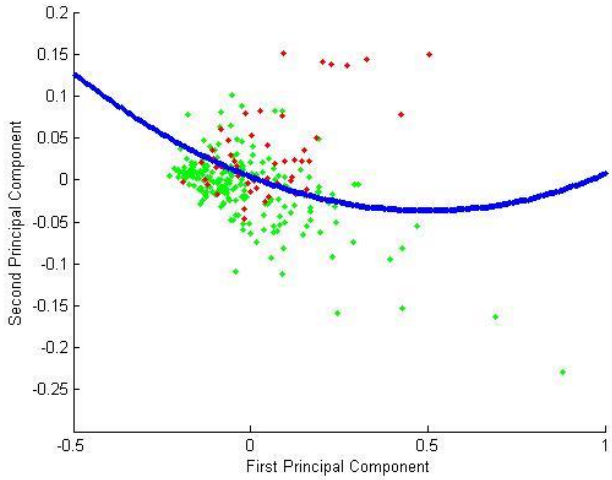

Fig. 4 An example robust solution produced by the established radial basis function based support vector machine ($\gamma = e^{-2}$).

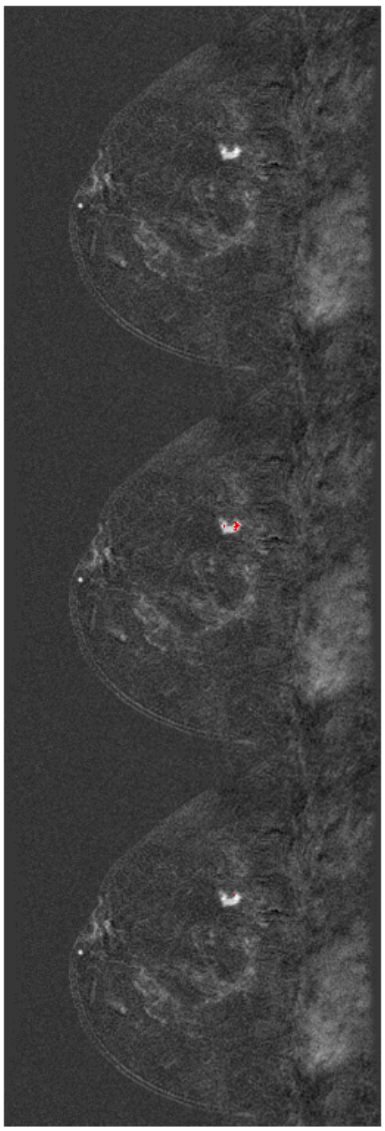

Figure 5. A breast MRI examination exhibiting a malignant lesion (top pane), the computer-aided diagnosis results of a robust support vector machine ($\gamma = e^{-2}$, middle pane) and computer-aided diagnosis results of an overfitted support vector machine ($\gamma = e^{12}$, bottom pane). Both classifiers were trained on the same data and have the same bias setting.

DISCUSSION

In this research study we presented a plotting technique for identifying overfitted classification functions. The technique presented can be useful as a simple check to make sure that the classification researcher is not producing overfitted solutions, but can also be generally used to evaluate classifier behaviour and to ensure that solutions produced by standard validation methods conform to a robust classification solution. Overfitted solutions typically underperform when tested on an independent dataset. Furthermore, we showed in figure 5 (bottom pane) that an overfitted solution can also be incapable of identifying as much of a malignant lesion when compared with a robust classification solution (middle pane). This can occur when the classification researcher trains the algorithm on MRI tissue sample measurements from the most hyperintense region of the tumour. In these circumstances a non-robust overfitted solution may be able to identify the most malignant appearing tissues available, however, the lack of function robustness means the algorithm has trouble identifying less obvious sections of the malignant tumour. Note that the overfitted support vector machine classifier was only able to identify a single voxel as malignant, whereas the robust support vector machine was able to identify two separate hyper-intense regions of this malignant lesion (correctly identifying 15 times as many malignant voxels as the overfitted support vector machine). Figure 5 represents a simple example of the potential consequences of using overfitted classification functions.

When examining the plots presented in this paper, overfitting can easily be qualitatively identified as a blue boundary whose border fluctuates greatly across our data space. This often occurs due to the decision function definition being extremely closely associated with one of the two classes of training data. For instance, in figure 2 the support vector machine produces an overfitted solution closely reflecting the distribution of malignant (red) samples. Similarly, in figure 3 the k-nn density classifier produces an overfitted solution closely reflecting the distribution of malignant (red) samples.

The established support vector machine based on the radial basis function can produce overfitted solutions when the γ parameter is set too high. The k-nn density classifier can also produce overfitted solutions but it has a tendency to do this at low *k* settings. Both of these classification techniques are able to produce overfitted solutions based on the same underlying principle. At high γ settings and at low *k* settings both the classifiers presented have been heavily biased towards having prediction based on a few of the samples that are closest to the test sample. Over-emphasizing this effect with any classifier will lead to an overfitted solution. Figure 4 presents robust classification solutions based on the support vector machine method. Although a wide range of parameter settings for the k-nn density classifier were experimented with, none of the *k* parameter settings yielded solutions that would qualitatively be described as being as robust as the support vector machine in figure 4.

We wanted to include comparative plots demonstrating that the popular artificial neural network can also produce overfitted solutions. However, there are an infinite number of possible artificial neural network architectures and implementations available to the researcher. Unfortunately there is no accepted gold-standard implementation of an artificial neural network for breast cancer diagnosis from magnetic resonance imaging examinations. As such future work will look at a detailed analysis of the artificial neural network by investigating its generalizability and its ability and tendency to produce overfitted solutions. This will involve comparing a variety of artificial neural network architectures and publicly available implementations of the technique.

The plotting technique demonstrated in this study is limited by the projection method selected. Principal components analysis provides the researcher with knowledge of how much variance is represented in the final plot. The plots presented in the results of this research paper represent 98% of the total variance in the dataset indicating that they are reliable plots on which to form conclusions regarding whether a solution is overfitted. If the percentage of total data variance in the final projected plot is deemed inadequate to perform a reliable comparison, then a set of pair-wise principal component plots can be created (ex. first principal component plotted against the second principal component as well as the second principal component plotted against the third principal component etc.). Future work will look at applying different data projection techniques for the creation of the initial two-dimensional plot. This technique is limited to those projection methods that can perform a reverse rotation from the plot's two-dimensional space back into the input space.

When standard validation techniques (k-fold cross validation, bootstrapping, etc.) produce two solutions with very similar performance metrics, plotting techniques like the one presented in this study can also assist the researcher in making a final selection of which classifier is likely to be the most robust.

Identifying overfitted classification solutions through classifier visualization is an evaluative method not widely used in the supervised learning literature. Although the plotting technique presented will not be appropriate in all applications, it is hoped that the

provided results will act as an impetus for researchers to evaluate their supervised learning studies through some type of classifier visualization. Classifier visualization has the potential to improve the quality of the results of supervised learning based pattern recognition research.

CONCLUSIONS

We have demonstrated that the presented classifier visualization technique can be an effective method for qualitatively identifying overfitted classification solutions and demonstrated the technique's potential benefit on breast cancer lesion data from magnetic resonance imaging examinations. The technique is recommended for use as an adjunct evaluation method to be combined with standard statistical validation techniques. The method presented can help the researcher avoid overfitted solutions while confirming robust classification results.

ACKNOWLEDGEMENTS

We would like to thank Gary Wang for his assistance in image registration and Elizabeth Ramsay and Don Plewes for their assistance in image acquisition. The magnetic resonance imaging data was provided by Dr. Ellen Warner and was acquired using funding from the Canadian Breast Cancer Research Alliance. The authors would like to thank the Canadian Breast Cancer Foundation and the Canadian Institute for Health Research for contributing to funding this research study.

REFERENCES

[1] Levman J, *et al.* Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines. *IEEE Transactions on Medical Imaging* 2008; 27(5):688-696.

[2] Efron B, *et al.*, *An Introduction to the Bootstrap*. New York, Chapman & Hall, 1994.

[3] Eng J. Receiver Operating Characteristic Analysis: A Primer. *Academic Radiology* 2005; 12:909-916.

[4] Warren R, *et al.* A test of performance of breast MRI interpretation in a multicentre screening study. *Magn Reson Imaging* 2006 24(7):917-929.

[5] Levman J, *et al.* Effect of the Enhancement Threshold on the Computer-Aided Detection of Breast Cancer using MRI. *Academic Radiology* 2009; 16(9):1064-1069.

[6] Nattkemper TW, *et al*. Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods. *Artificial Intelligence in Medicine* 2005 34:129-139.

[7] Nattkemper TW, Wismuller A. Tumor feature visualization with unsupervised learning. *Medical Image Analysis* 2005 9:344-351.

[8] Komura D, Nakamura H, Tsutsumi S, Aburatani H, Ihara S. Multidimensional support vector machines for visualization of gene expression data. *Bioinformatics* 2005 21(4):439-444.

[9] Somorjai R, Dolenko B. Class-preserving mapping of high-dimensional biomedical data: visualization, classification, clustering. *Proceedings Interface and Classification Society of North America*, St. Louis, MO, USA, June 2005.

[10] Hawkins DM. The Problem of Overfitting. *J. Chem. Inf. Comput. Sci.* 2004 44(1):1-12.

[11] Babyak MA. What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine* 2004 66:411-421.

[12] Schaffer C. Overfitting avoidance as bias. *Machine Learning* 1993 10(2):153-178.

[13] Reunanen J. Overfitting in making comparisons between variable selection methods. *The Journal of Machine Learning Research* 2003 3:1371-1382.

[14] Tetko IV, Livingstone DJ, Luik AI. Neural network studies. 1. Comparison of overfitting and overtraining. *J. Chem. Inf. Comput. Sci.* 1995 35(5):826-833.

[15] Larimore WE, Mehra RK. Problem of overfitting data. *Byte* 1985 10(10):167-178.

[16] Lawrence S, Giles CL, Tsoi AC. Lessons in Neural Network Training: Overfitting May be Harder than Expected. *Proceedings of the Fourteenth National Conference on Artificial Intelligence* 1997, Menlo Park, California, 540–545.

[17] Lampinen J, Kostiainen T. Self-Organizing Map in Data Analysis – Notes on Overfitting and Overinterpretation. *European Symposium on Artificial Neural Networks* 2000, Brussels, Belgium 239-244.

[18] Gianniotis N, Tino P. Visualization of Tree-Structured Data Through Generative Topographic Mapping. *IEEE Transactions on Neural Networks* 2008 19(8):1468-1493.

[19] Weigend AS, Mangeas M, Srivastava AN. Nonlinear gated experts for time series: discovering regimes and avoiding overfitting. *International Journal of Neural Systems* 1995 6:373-399.

[20] Karystinos GN, Pados DA. On Overfitting, Generalization, and Randomly Expanded Training Sets. *IEEE Transactions on Neural Networks* 2000 11(5):1050-1057.

[21] Monari G, Dreyfus G. Local Overfitting Control via Leverages. *Neural Computation* 2002 14(6):1481-1506.

[22] Sarle WS. Stopped Training and Other Remedies for Overfitting. *Proceedings of the 27th Sumposium on the Interface*, 1995.

[23] Lee Y, Lee S, Ivrissimtis I, Seidel HP. Overfitting Control for Surface Reconstruction. *Eurographics Symposium on Geometry Processing*, 2006.

[24] Jacobs MA, *et al*. Benign and malignant breast lesions: diagnosis with multiparametric MR imaging. *Radiology*, 2003 229:225-232.

[25] Gilhuijs KGA, Giger ML, Bick U. Computerized analysis of breast lesions in three dimensions using dynamic magnetic-resonance imaging. *Medical Physics* 1998 25:1647-1654.

[26] Chen W, Giger ML, Lan L, Bick U. Computerized interpretation of breast MRI: Investigation of enhancement-variance dynamics. *Medical Physics* 2004 31(5):1076-1082.

[27] Chang CC, Lin CJ. LIBSVM [Online] (Accessed Jan. 25, 2009) http://www.csie.ntu.edu.tw/~cjlin/libsvm/